German Working Papers in Law and Economics

Volume 2003 *Paper* 3

A Difficulty with Oaths: On Trust, Trustworthiness, and Signalling

Friedel Bolle
Europa-Universität Viadrina

Matthew Braham
University of Hamburg

Abstract

In the wake of the Enron and Worldcom financial scandals that rocked Wall Street in 2002, the US government's financial regulatory body, the Security and Exchange Commission (SEC) took the unprecedented step in June 2002 of requiring that the chief executives and chief financial officers of America's 947 biggest companies to swear on oath that their company results and financial reports were to the best of their knowledge accurate. The one-off order was quickly followed by the passing of the Sarbanes-Oxely act, which will require many more CEOs and CFOs to certify their company reports and financial statements at regular intervals. In this paper we apply a simple signalling model to examine whether or not this type of institutional signal of trustworthiness is always efficient. We find that in the presence of signalling costs, the separating equilibrium can be socially inefficient as well as causing a general loss of trust.

A Difficulty with Oaths: On Trust, Trustworthiness, and Signalling

Friedel Bolle
Faculty of Economics
Europa-Universität Viadrina (Oder)
Postfach 1786
15207 Frankfurt (Oder)
Germany
bolle@euv-frankfurt-o.de

Matthew Braham
Institute of SocioEconomics (IAW)
University of Hamburg
Von-Melle-Park 5
20146 Hamburg
Germany
braham@econ.uni-hamburg.de

27 February 2003

Abstract In the wake of the Enron and Worldcom financial scandals that rocked Wall Street in 2002, the US government's financial regulatory body, the Security and Exchange Commission (SEC) took the unprecedented step in June 2002 of requiring that the chief executives and chief financial officers of America's 947 biggest companies to swear on oath that their company results and financial reports were to the best of their knowledge accurate. The one-off order was quickly followed by the passing of the Sarbanes-Oxely act, which will require many more CEOs and CFOs to certify their company reports and financial statements at regular intervals. In this paper we apply a simple signalling model to examine whether or not this type of institutional signal of trustworthiness is always efficient. We find that in the presence of signalling costs, the separating equilibrium can be socially inefficient as well as causing a general loss of trust.

Keywords asymmetric information, institutional signals, oaths, risk, trust

JEL classifications C72, D81, D82, K22

A Difficulty with Oaths: On Trust, Trustworthiness, and Signalling*

Friedel Bolle and Matthew Braham

… I say unto you, Swear not at all; …
— Jesus in the *Sermon on the Mount* (Matthew 5, 34)

I swear … that to the best of my knowledge (which is pretty poor and may be revised in future), my company's accounts are (more or less) accurate. I have checked this with my auditors and directors who (I pay) to agree with me.
— *The Economist*, 17 August 2002.

1. Introduction

In the wake of the Enron and Worldcom financial scandals that rocked Wall Street in 2002, the US government's financial regulatory body, the Security and Exchange Commission (SEC) took the unprecedented step in June 2002 of requiring that the chief executives (CEO) and chief financial officers (CFO) of America's 947 biggest companies[1] to swear in front of a notary that 'to the best of my knowledge', their latest annual and quarterly reports neither contain an 'untrue statement' nor omit any 'material fact' relevant to investors. Although the order was a one-off, it was quickly followed by the passing of the Sarbanes-Oxely act which was rushed into law just over a month later. This law will require many more CEOs and CFOs to certify at regular intervals that their reports and financial statements contain no untruths or omit material facts.[2]

* We would like to thank Winand Emons and other participants at the Hamburg Law and Economics Workshop, 14 February 2003, for comments.

[1] Companies with annual revenues of more than $1.2 billion.

[2] For background, see *The Economist* (17.08.02) and www.cfo.com: 'I Solemnly Swear …' (05.08.02) and 'Is Signing Certifiable' (05.08.02).

The move by the SEC and US law makers was obviously an attempt reduce volatility in the financial markets by finding a means to reassure an increasingly sceptical public that company bosses are an honest bunch. Not all of them were extracting $6,000 gold-and-burgundy floral-patterned shower curtains on their expense accounts as one executive was doing, or like the CEOs of Enron and Worldcom claiming that they had no idea that their staff were fiddling the books. By issuing an order to swear on oath and regularly certify their accounts the regulatory authority and the law maker was introducing a powerful incentive to be honest. Perjury, or lying under oath, is a criminal offence with severe penalties. Some commentators have speculated that CEOs and CFOs who falsely certify their company's financials could, in the most egregious of case, be faced by fines of up to $5 million and 20 years jail.[3]

The SEC's one-off ruling and the Sarbanes-Oxely act raise two questions. The first is obviously whether or not swearing oaths will actually make company bosses tell the truth and prevent exploitation of investors; the second is whether it is a socially desirable mechanism even if it could elicit the truth. This paper deals with the second question; but in doing so we presume (maybe wrongly) that the first question is answered positively: we take it that perjury is an effective incentive to tell the truth if one decides to swear on oath.

To answer the second question it is necessary to turn it around a little and ask, what value is the truth that is to be elicited by swearing on oath? In this particular case, the aim seems to be one of increasing the amount of trust that investors have in corporate bosses – trust that these bosses will not bankrupt the company and embezzle its funds by hiding expense accounts for shower curtains. And the value of trust is, of course, that it reduces transaction and monitoring costs. As Arrow (1974: 23) remarked, 'Trust is an important lubricant of the social system'.[4]

Thus we can ask, will such an order for CEOs and CFOs to certify under oath as to the veracity of their company reports and financial statements increase in the aggregate the amount of trust in the economy and lead to an increase in social welfare?

---

[3] 'I Solemnly Swear …' (05.08.02), www.cf.com

[4] Note that the value of trust for the economy was recognized by Adam Smith (1776 [1981]: 66) in *The Wealth of Nations*, where he noted that trust is rewarded in wages.

We answer the question by recourse to a very simple signalling model. We take it that it is fairly self-evident that an oath is a quality signal, although one of a special kind, which we call an *institutionalized signal*. This is a signal which is enforced by law or convention. It has the characteristic that non-compliance with an order to send the signal can also convey important information to the recipient. For, not swearing on oath is a signal itself and will, under certain circumstances, be interpreted negatively by its recipients: in this case investors who will take it to indicate that something is awry with the company and therefore pull their capital out of the company before it is too late. It appears that this is precisely what the SEC and the Sarbanes-Oxley act are trying to achieve: a separating equilibrium in which honest bosses swear on oath while (as a simplification) dishonest ones do not (or at least have to restate its results before doing so[5]). That is, sworn certification is a sorting mechanism for investors to determine which companies are taking their responsibilities as a fiduciary seriously and are acting in the interests of shareholders. The result that we arrive at is that this form of institutionalized signalling can have a twofold effect that run contrary to its apparent *raison d'être*: it can reduce both the amount of trust and social welfare.[6]

2. Socially Harmful Signalling

2.1 *The Basic Model*

Our model is a simple three stage game in which a firm and investor meet and decide on an investment contract. In order to demonstrate our main thesis, that institutionalized signals such as requiring the CEO or CFO or a firm to swear on oath as to the veracity of the firm's book-keeping can reduce trust and social welfare we keep our model as simple as possible and allow asymmetric information into the model only in two minimal respects. This will be sufficient to produce our result of the existence of an inefficient separating equilibrium.

---

[5] In the model that follows, we take restatement of company results prior to swearing on oath to their veracity as not sending the signal.

[6] This result differs from the literature in that in most signalling games such as Spence (1973, 1974), Camerer (1988), Carmichael and MacLeod (1997), and Bolle (2001), and others the resulting equilibria are investigated, but not the question of whether the introduction of the signal is beneficial.

Our game has two players: a firm (F) and an investor (I). The firms are of two principle types: they are either honest or exploit. We assume that there is only one type of investor. This is simplifying assumption, as it says that each firm faces a population of investors with the same payoff function, i.e. its payoff's (u_I and y_I) depend on the firm and not his or her own type. We do not believe this assumption to be too severe in the context of the question we seek to answer.

We also assume that transaction costs make legal enforcement of the contract infeasible – this is the essential element in a trust relationship.[7] And finally, we make the usual assumption of players having von Neumann-Morgenstern utility functions and being risk neutral, i.e. they want to maximize expected income.

The difference between the types of firms is characterized by the payoff. Honest firms are precisely those whose payoff from holding to the terms of the investment contract is greater than that of breaching the contract and embezzling the surplus at the expense of the investor. If this game of trust is one of full information, the equilibrium outcome is trivial: only matches between investors and honest firms come into being. Signalling is naturally irrelevant here. Where it becomes relevant is of course where the game is one of incomplete information. We introduce this in a minimal way by making the firm's type (what its payoff happens to be) private information. All that the investor knows is the distribution of types, and not whether a particular firm is trustworthy; put another way, from the investor's point of view, the firm's profits and costs functions are drawn from a certain distribution.

The introduction of an institutional signal changes the game of trust in one very important respect. If the signal is sufficiently costly (which is the main feature of our model), which we assume to be the case, in that the probability of detection and costs punishment for imitating the signal (making a false oath) are sufficiently high as to *always* deter an exploiter (type E) from sending the signal, then investors will usually prefer to make a match with a firm that has sent the

[7] Here we follow, in particular, Coleman (1991: 91) in taking situations involving 'trust' to refer to constitute a sub-class of those involving risk; situations in which the risk one takes depends on the performance of another actor. That is, trust is essentially a game theoretic and not a classical decision-theoretic problem. See also Coleman (1984). For a discussion of this notion of 'trust as risk', see Dasgupta (1988) and Williamson (1993) and the references therein. For more general background, see the contributions in Gambetta (1988).

signal than with one that has not. That is, if a signal is an unfakeable cue, a strategic problem is now imposed on the firms: to invest in the signal or not.

Thus once a signal is instituted, an honest firm can be honest and wealthy (type H) – it can afford the signal; or honest and poor (type $\tilde{H}$) – it cannot afford the signal.[8] The latter case is one in which expected profit of a match with an investor are now offset by the cost of the signal to such an extent that the outside option of zero, which we assume results if an investor does not decide on a match, is preferable to sending the signal and obtaining a match. Thus, we have a situation in which firms can discern only honest and wealthy firms (type H) but not between honest and poor (type $\tilde{H}$) and the exploiters (type E). However, whether or not a firm can afford the signal has no effect upon the investor's payoff; it only affects the firm's; an investor receives the same payoff for trusting an honest firm regardless of the firm's ability to cover the cost of the signal.

The structure of the game is summarized by Figure 1. The sequence of play and the payoffs are as follows:

Stage 1 Nature randomly determines the type of firm on the basis of a continuous, strictly increasing distribution function, i.e. a firm's u_F, y_F, and c_F are random variables. Where u and y designate payoffs and c the cost of the signal (subscripts F and I refer to firms and investors respectively). Firms know their type; investors do not know it, although they know the proportion of honest and exploiting firms in the total pool of firms.

Stage 2 Firms decide to send a signal (S) or not (NS).

Stage 3 The selected type of firm and the investor meet. The investor determines whether or not to trust (T) the firm. If she does not trust (NT), both receive a payoff of zero and the game is terminated. If she trusts, the game proceeds to stage 3.

Stage 4 The firm decides to cooperate and fulfil (C) the terms of the investment contract or not (NC). If the firm is honest, the investor receives a payoff of u_I and

[8] Note that the attributions 'wealthy' and 'poor' are not to be taken literally. They only mean that from the firm's point of view it is optimal or suboptimal to send the signal. This has to do with the specific profits and costs that are involved in the firm's (the executive's) decision to send the signal or not.

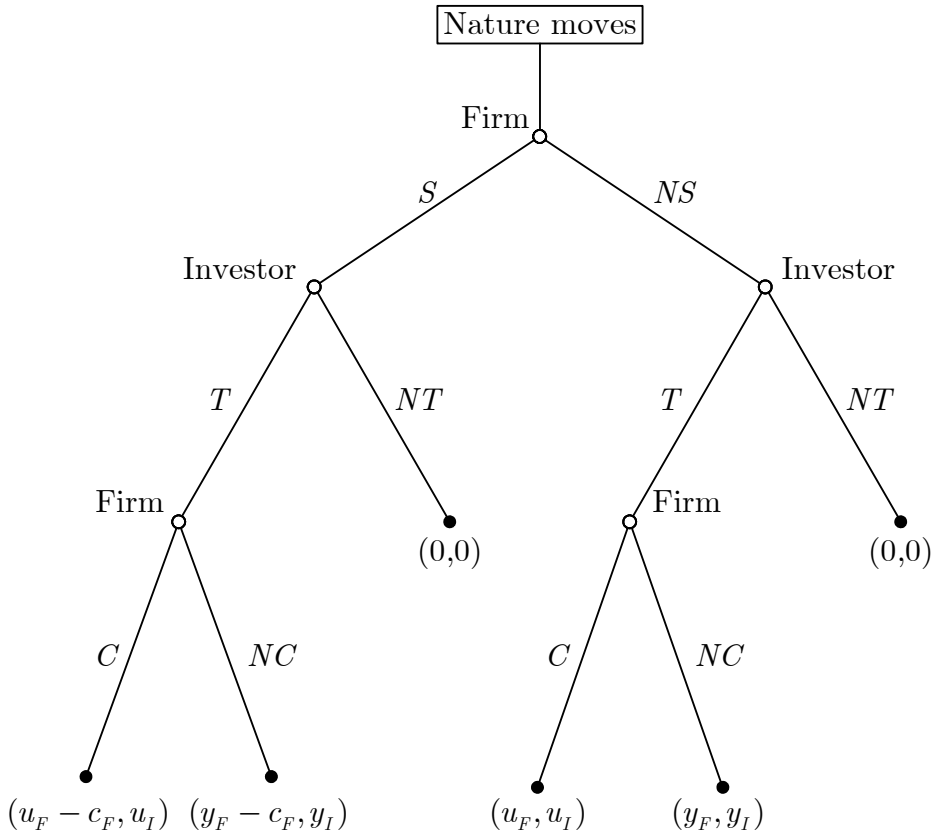

Figure 1

the firm a payoff of u_F; if the firm is an exploiter, it breaches the contract and embezzles the investor's money giving the investor a payoff y_I and himself a payoff of y_F, with $y_F > u_F$. We assume that $u_I, u_F, y_F > 0$, $y_F < c_F$ and $y_I < 0$. We further assume that firms' profits and costs are private knowledge and that the investors profits are constant and common knowledge. Firms with $u_F > y_F$ are 'honest' (type H or $\tilde{H}$); firms with $u_F < y_F$ are 'exploiters' (type E).

2.2 *No-oath Equilibria*

As we have explained, we want to show that the institutionalization of a signal can, paradoxically, decrease trust and cause a welfare loss even though the proportion of trustworthy firms remains unchanged. To do this, we first consider the case of there being no institutionalized signal, i.e. firms are neither required by law or convention (social norms) to swear on oath as to the veracity of their book

keeping nor do they have the possibility to do so voluntarily by some means.[9] This is obviously an extreme assumption and does not reflect the reality of any legal order. But we consider this as our baseline scenario.

If in this baseline scenario investors trust firms, the situation is the same as a *pooling equilibrium* in the signalling game of Figure 1 in which *no* firm sends a signal and *all* investors trust and invest or *no* investor trusts (and makes do with the outside option of zero). Let V be the expected value of an investor's trust without a signal:

$$V = \left(P(H) + P(\tilde{H})\right) \cdot u_I + P(E) \cdot y_I \tag{1}$$

With $P(H) = \Pr(u_F > c_F)$, $P(\tilde{H}) = \Pr(c_F > u_F > y_F)$, $P(E) = \Pr(u_F < y_F)$.

(Remember the generally true relation $y_F < c_F$ which is also known by the investors.)

Proposition 1 (No Signalling) *If no signalling is possible and $V > 0$, then all investors trust (T) and no firm sends a signal (NS); if $V < 0$, then no investor trusts (NT). If signalling is possible and $V > 0$, then sending no signal (NS) and trusting (T) is a pooling equilibrium; if $V < 0$ such a pooling equilibrium does not exist.*

(Without proof.)

Remark 1 Only for certain parameter values will (1) be an equality. If these parameters are randomly chosen, then $V = 0$ only with a probability zero. We do not consider this (non-generic) case further.

In words, investors will trust – take a risk – if either there are a sufficient number of trustworthy firms in the population, or the payoff from investing in a trustworthy firm is large enough even though there is a positive probability of being exploited.

---

[9] We assume that conventions can be as authoritative as formal law. This assumption is in line with Basu's (Basu 2000: 117) 'core theorem of law and economics': 'Whatever behavior and outcomes in society are legally enforceable are also enforceable through social norms'.

Given that we ultimately wish to compare two institutional arrangements, one without signalling and one with, we need compare the social value of each arrangement. For this we assume a simple utilitarian social welfare function, W (the case of all firm types being trusted without signals) given by:

$$W = P(H)(u_I + u_F^H) + P(\tilde{H})(u_I + u_F^{\tilde{H}}) + P(E)(y_I + y_F^E) \quad (2)$$

Note that in the case of the pooling equilibrium described by Proposition 1, u_F^H and $u_F^{\tilde{H}}$ and c_F are mean values as defined by H, $\tilde{H}$, and E. The other payoffs are assumed by the model setup to be constant. We will make use of (2) in section 2.3 below.

2.3 *Oath Equilibria*

Our next step is to investigate the conditions and implications of sending an institutionalized signal such as an oath attesting to the veracity of a firms accounts. The idea of this instrument is obviously to obtain a separating equilibrium in which all honest firms send the signal and the dishonest ones do not. If they can, all investors will make a match with honest firms and social welfare can increase because the losses accruing to investors being exploited will now be eradicated. It is obvious that we are almost guaranteed an efficient outcome if all honest firms *can* in fact send the signal (it is not too costly to do so) and the exploiters *can never* send the signal (it is too costly). Put another way, if the signal is cheap but unfakeable because the consequences of detection are too severe,[10] the introduction of an institutionalised signal will have the desired effect.[11]

However, matters change significantly in the presence of substantial signalling costs, c_F (with the addition of a H or $\tilde{H}$ superscript where relevant). The

---

[10] It could be disputed that unfakeable (or unimitable) signals cannot be cheap; if they were they could not be unfakeable. In particular this is a position taken with regard to animal signalling in evolutionary biology, known as the 'handicap principle' (Zahavi 1975, Krebs and Dawkins 1984). However, in our context it would seem possible. Consider the one-man firm. He knows with certainty whether or not he has fiddled the figures or not and betrayed the confidence of his investors. Certifying his accounts will be relatively cheap (essentially only the costs of the notary services). In contrast, a huge multinational corporation the same act – certifying on oath – will be extremely costly as it requires an elaborate accounting, reporting, and monitoring mechanism.

[11] If no investor would trust without a signal, it is certainly efficient to institutionalize a one.

signalling cost in the particular case that we are examining here is not only the actual outlay on procuring the signal, e.g. additional auditing and legal costs in the case of swearing that the firm has not engaged in 'erroneous book keeping' plus the opportunity cost of executives having to scrutinize company reports,[12] but also the cost of a false accusation of lying on oath. Perjury is a severe criminal offence in most countries. As we already said in the introduction, the punishment in the US for a corporate boss who falsely swears to the accuracy of his company accounts is estimated to be in the region of $5 million and 20 years imprisonment in the most severe of cases. It is vital to note that in our model we are not merely concerned with the expected loss due to the (in reality) rather small probability of a judicial error which ends up wrongly convicting a senior executive for perjury, but also the expected loss derived from an unsubstantiated accusation which may severely undermine a high ranking executive's reputation for trustworthiness and hence his or her future stream of income, a probability that in reality may not be negligible. So long as investors are aware that a lack of proof of embezzlement is not necessarily evidence of being honest, then being accused of dishonesty may have serious repercussions even though one is in fact honest. It comes as no surprise to us that given the seriousness of perjury, legal experts were advising executives not to make the oaths (particularly as it was still unclear what the legal sanctions would be for not doing so).[13] The point being that if the financial market cognoscenti take the missing paperwork as a sign that something may be amiss with a company and hammer it on the belief that it has something to hide, for the honest executive this may still be better than being falsely accused of lying under oath (even without it leading to a false conviction).

Now, the separating equilibrium can be inefficient compared to the no oath arrangement as the honest group for whom the cost is too high will not send the signal and therefore will not necessarily find a match with an investor. To see that

---

[12] For example, shortly after the SEC issued the order in June, the CEO of Sun Microsystems, Scott McNealy, complained that compliance would force him to spend valuable time scrutinizing financial statements instead of drumming up business by visiting customers. See, 'Is Signing Certifiable', www.cfo.com (05.08.02).

[13] See, 'Is Signing Certifiable', www.cfo.com (05.08.02). Consider the risk that a chief executive faces if the Economist's caricature of oath-taking which we gave in the epigraph of this essay would not be so far from reality. The lack of knowledge that a senior executive has indicates just how costly swearing on oath may be.

this is equilibrium behaviour (for the investor to play *NT*), an investor will *not* trust in absence of a signal if the expected value of trust V' is:

$$V' = \frac{P(\tilde{H})}{P(\tilde{H}) + P(E)} \cdot u_I + \frac{P(E)}{P(\tilde{H}) + P(E)} \cdot y_I < 0 \tag{3}$$

We now can state the following simple proposition:

Proposition 2 (Signalling) *If, for all* y_F^E, $y_F^E < c_F^E$, *and defining the probabilities in (3) as in (1), then we have three cases:*

Case I *If (1) and (3) hold then we have two equilibria: (a) a pooling equilibrium in which all investors trust (T) and no firm signals (NS) – do not sign the oath; and (b) a separating equilibrium in which investors only trust (T) those firms which send the signal so that only type H firms signal (S) – sign the oath – and types* $\tilde{H}$ *and E do not send the signal (NS).*

Case II *If only (1) holds we have a pooling equilibrium (equivalent to Proposition 1).*

Case III *If only (3) holds we have a separating equilibrium (as in Case I(b)).*

Proof It is easily checked that given the behaviour of others, it is optimal to stick with the behaviour described in each case in the proposition. Note that $V' > 0$ implies $V > 0$ so that there is no fourth case. If $P(H) = 0$, then the separating equilibrium is intrinsically a pooling equilibrium without signalling (*NS*) and without trust (*NT*). ■

Remark 2 In Case I there is also a third, mixed strategy equilibrium (Bolle 2002), but we do not investigate it here. Also note, that for our purposes, Case I is the interesting case. Although from a game theoretic point of view we have two equilibria, from a behavioural point of view we are likely only to have the pooling equilibrium because it does not pay for a single firm to send the signal. That is, the separating equilibrium can only be instituted from the outside.

Now, in order to demonstrate our claim oaths can be inefficient, we define social welfare under institutional signals, W', as:

$$W' = P(H)(u_I + u_F^H - c_F^H) \tag{4}$$

The relative efficiency of the two systems can be assessed by the differences in welfare, Δ, which is given by:

$$\begin{aligned} \Delta &= W - W' \\ &= P(\tilde{H})(u_I + u_F^{\tilde{H}}) + P(E)(y_I + y_F^E) + P(H)c^H \end{aligned} \tag{5}$$

The following proposition is straightforward:

Proposition 3 (Inefficiency) *There exist separating signalling equilibria that are inefficient with respect to the no signalling pooling equilibrium.*

Proof By example. Assume $P(H) = P(H) = P(E) = \frac{1}{3}$ and (1), (3), and the conditions for Case I of Proposition 2 hold. Assume additionally, $c_F^H = c_F^{\tilde{H}} = c_F^E = C > \frac{1}{3}$; $u_I = 1$, $-2 < y_I < -1$; $u_F^H = 1$, $u_F^{\tilde{H}} = \frac{1}{3}$, $y_F^E = \frac{1}{3}$. Plugging Plugging these values into (5), we obtain $\Delta > 0$ for $y_F > -\frac{5}{3} - c_F^H$.[14] ■

Remark 3 Note that for $-\frac{5}{3} < y_I < -1$ the separating equilibrium is inefficient even in the absence of signalling costs c_F^H. What this implies, then, is that signalling costs are important for efficiency, but not decisive. The social product can be sufficiently reduced by a deterioration in trust alone (trustworthy firms nit finding a match). Note, therefore, that the result is not driven by the size but only the function of the signalling costs.

Intuitively, if the signal is unaffordable for some firms, then the effect of the signal is to raise the conditional probability that an unmatched investor will end up with an exploiter once all firms that could afford the signal have found a match. This can lead to a social loss even without taking into account the price of the signal for those who could afford it. For a more light-hearted example, consider the case of six men and six women. Men can be 'husbands' who want a long-term relationship ('until death do us part') or 'Latin lovers' ('here today, gone tomorrow'). Women are 'wives': they prefer husbands to Latin lovers. Imagine now that husbands can be distinguished by gifts of diamond engagement rings ('diamonds are woman's best friend). Latin lovers would never make such

[14] Note that the bounds on y_I are required to satisfy (1).

an investment (far too expensive for fleeting romance). Now assume further that there are four 'husbands' and two Latin lovers, but only two husbands can afford a diamond ring; call those who cannot afford diamonds 'bachelors' (i.e. willing but unsuccessful 'husbands'). If the convention is to give diamond engagement rings, then there are only two matches; the two other men with good intentions and the two other women who want a husband go without a pairing, although both couples would prefer this. However, without the convention of diamond engagement rings, we would have four matches: all well intentioned men (husbands + bachelors) who want a wife get one, even though two unlucky women have to suffer the consequences the morning after … On the whole society will be better off *without* the institutionalized signal of diamond engagement rings (and possibly engagement *tout à fait* if it is an expensive custom) as long as the utility losses of the two women who end up with Latin lovers does not exceed the collective gains of the four pairs and the two men they end up with.[15]

The interesting feature of this marriage-market example is that it shows there need not be a social loss involved in the act of signalling itself even if it reduces trust: giving a diamond ring only means a transfer of the social loss.[16]

3. Discussion

The basic result of this paper says that for two societies which are alike in all respects bar the presence of an institutionalized signal in one but not the other, will have different levels of social welfare as measured by the social product (W). The importance is that they both have the equivalent amount of trustworthy individuals, but not the equivalent amount of trust, i.e. individuals who are willing to take a risk that another will actually perform what they have agreed to in circumstances where the agreement cannot easily be enforced by legal means. This can be sufficient to lower the social product even before taking into account the reduction brought about by those who have actually invested in the cost of the signal. To give more of a paradoxical flavour to the result, our model also shows

[15] This marriage game is analysed in more detail in Bolle (2001). See also Camerer (1988)

[16] Assuming, of course, a model with only 'wives'; there is no 'reverse exploitation' in which 'virigins' collect diamond rings.

that it is quite possible that one society has more trustworthy individuals than another ($P(H_1)+P(\tilde{H}_1)>P(H_2)+P(\tilde{H}_2)$) but has less trust and a lower social product. Whether this will occur obviously depends on how large $P(H)+P(\tilde{H})$ is in general, and more specifically the value of $P(\tilde{H})$. If trust is – in Arrow's words again – 'an important lubricant of the social system', then instituting a signalling mechanism can be the spanner in the works![17]

To phrase the result in another way, relations of trust have two components: a trustee and a trustor. What our model shows is that being a trustworthy trustee (a trustee upon whom a trustor can rely to implement the terms of an agreement without external sanctions) is not a sufficient condition to be trusted. The placement of trust is affected by the probability of encountering a trustworthy trustee in a random matching and by the institutional environment governing signals of trustworthiness. These signals are costly, with the result that if instituted in an environment with a sufficiently large proportion of trustworthy individuals who cannot afford the signal, then the signal can be counter productive. Thus, to take the case which motivated this paper, the SEC ruling that CEO's and CFO's swear on oath that their business accounts are accurate will certainly increase reliability and trust of those who give the oath; but there may be a large number of managers who, in principle, are trustworthy but who are not ready to take the risk of the oath. As not signing the oath is a negative signal under these circumstances, not only corrupt but also reliable companies will suffer a loss of trust. In aggregate this may or may not increase efficiency (the social product). In a nutshell, the aggregate amount of trust in an economy is *not* monotone in the aggregate amount of trustworthiness.

One of the interesting features of our result is that it is applicable to a wide range of phenomena. Institutionalized signalling is as much a matter of social conventions as legal rules and regulations. Our marriage market example is a case in point. Another example is the job market. In countries such as Germany, one can only set up a business as a baker, butcher, carpenter, electrician, etc., if one has a *Meisterbrief*, which is a certificate of 'Master Craftsmanship'. This is clearly a quality signal in the sense of Spence's (1973, 1974) classic study.

[17] It is interesting to note that in a very simple experimental set up of a marriage market based on this theoretical set up of this paper, Bolle and Kaehler (2002) have been able to confirm that that behaviour is fairly consistent with our model so that the use of gifts as a signal of being a honest can reduce trust ('marriages') and the social product.

Although it is a basic feature of Spence's model that signalling (separating) equilibria could be Pareto inferior, his reason was different to ours. In his model, the inefficiency arises when one tries to increase the level of education and training (the signal) above a threshold amount such that it does not improve the quality of sorting by applicants by employers by one bit. In our model, the inefficiency arises if there are able employees who cannot afford the signal; or in the case of Germany, able craftsmen who cannot afford to purchase a training that leads to *Meisterbrief*. In this latter case not only do able craftsmen who are not permitted to setup business suffer, so to do the consumers as the supply of craftsmen and competition is reduced leading to a higher equilibrium price. The only ones who benefit are those privileged enough to be able to incur the sunk costs of their title. Consider the loss to the economy because able managers are unable to purchase an MBA from an appropriate business school. It does not take much of a stretch to the imagination to see that the model applies to a whole range of social practices related to the acquisition of status that determines entry into various job markets.[18]

In short, what we have shown is that while it is true that if a signal is a sufficiently expensive and unfakeable cue, it entails that the sender possesses a particular quality, it is *not* true that the absence of a signal entails non-possession of the quality. Clearly this is a logically trivial conclusion; but its implications for institutional design are not. So, for instance, when it is observed that 'among the properties of many societies whose economic development is backward is a lack of mutual trust' (Arrow 1974: 26) it should not be concluded that these societies *lack* the institutions and rules for signalling trustworthiness. The fact is, these societies may have *too* much signalling in the form of social customs and behavioural codes which prevent – in a mouthful – trustors trusting trustworthy trustees.

As a caveat, it could be argued that our result is not particularly important because an evolutionary argument along the lines of Alchian (1950) would mean that either inefficient signalling institutions would not emerge, or if, by chance, they did, they would not persist for very long as inefficient behaviour and institutions are driven out by competition. This is not the place to examine the

[18] In this regard, our results touch upon complementary studies of the effects of social custom. See for example Akerlof (1976, 1980) and Basu (1989).

problem in detail, but we would like to state our position on it. We believe that an evolutionary mechanism may not necessarily drive out an inefficient signalling institution. If the different parties have different interests it seems to us quite possible that one side of the match prefers the inefficient separating equilibrium. If we take our quaint marriage market example again, it is plausible to conceive that women as a group would be better off with signalling (diamond engagement rings) while the men suffer. Thus women (or investors) may try, and succeed in, enforcing the signalling institution, although this is harmful for society as a whole.

4. Concluding Remarks

To tie up this paper, two main qualifications are in order. Firstly, our model is rudimentary: the only feature which really matters in signalling is its cost. Clearly this is unrealistic in many a context. Thus to return to the motivation of this paper, we do not claim that either the one-off SEC order that top executives personally certify company accounts or the Sarbanes-Oxley act which will require certification at regular intervals is inefficient. Our claim is weaker: neither the one-off oath nor Sarbanes-Oxley can guarantee a welfare improvement, as measured by the social product, W. To make the other claim would require a much richer model that takes into account many more legal and institutional details.[19] We were content only with determining whether or not it is always socially desirable to introduce an institutional signal such as oaths. Our answer is negative.

Secondly, it must be remarked that we have treated institutional signals such as oaths as what can be termed as a signal *simpliciter*, that is, as an unmistakeable cue that serves to communicate information which can incite action or influence the behaviour of others. But this is only a partial understanding of what oaths are about. Oaths, in particular, are often far more than this: they can also be seen as preference changing signals. They can be seen as a commitment device which create the preference structure necessary for trustworthiness in advance of the decision to cooperate or exploit; call these ‘preference-shaping signals’. Clearly these signals have the characteristic of changing intrinsic and extrinsic costs and

[19] For example, the penalties for non-compliance. At the time of writing this was still unclear what this would be.

benefits. If an oath is costly, because breaking it is a criminal offence (law of perjury), can ruin our reputation, or merely because we will suffer a bad conscience because of the value system that we possess, it is a way of providing currency to our promises. In this context, if we have to take an oath – send an institutionalized signal – we have reasons for doing something in the future even though we may no longer want to do it. In this way an oath can be seen as a way of creating trust and predictability in interpersonal relations.[20] Thus if an executive knows that he has to certify his company's accounts on oath, the incentives that come along with the severity of perjury can change an executive preferences from one of acting negligently (not reading the company accounts) to acting with due care and attention (reading the company accounts). What was one onerous may now be 'pleasurable'. Clearly our result, that institutionalized signalling such as taking an oath, can easily be inefficient does not take into account this more subtle and complicated texture of signals.

Finally, despite the rough cut of the model we do believe that the wedge that the result drives between the 'amount of trustworthiness' and the 'amount of trust' in a society is important enough to earn further theoretical speculation and empirical and experimental investigation. It certainly does not seem to be an obvious fact that an institution designed to signal the integrity of trading partners could lead to the withdrawal of trust and thus to a reduction in society's welfare. This would appear to be particularly relevant to the institutional features of economic development and growth as well as to the more abstract study of the social organization of trust relations.

References

Akerlof, G.A. (1976), The Economics of Caste and the Rat Race and Other Woeful Tales, *Quarterly Journal of Economics* 90: 599–617.

Akerlof, G.A. (1980), A Theory of Social Custom, of Which Unemployment May be One Consequence, *Quarterly Journal of Economics* 94: 749–775.

Alchian, A.A. (1950), Uncertainty, Evolution, and Economic Theory, *Journal of Political Economy* 58: 211–221.

Arrow, K.J. (1974), *The Limits of Organization,* vol., New York: Norton.

[20] This is the standard position taken in the philosophy of law. See, for instance, Raz (1975: 69–70). For an economic approach to the question of preference-changing signals, see Bolle (1989).

Basu, K. (1989), A Theory of Association: Social Status, Prices and Markets, *Oxford Economic Papers* 41: 653–671.
Basu, K. (2000), *Prelude to Political Economy,* vol., Oxford: Oxford University Press.
Bolle, F. (1989), Levels of Aspiration Behavior, Promises, and the Possibility of Revaluation, in: K.W. Grunert and F. Ölsander (eds), *Understanding Economic Behavior*, Dordrecht: Kluwer, 199–212.
Bolle, F. (2001), Why Buy YOur Darling Flowers: On Cooperation and Exploitation, *Theory and Decision* 50: 1–28.
Bolle, F. (2002), Signals for Reliability: A Possibly Harmful Institution?, *Central European Journal of Operations Research* 10: 217–227.
Bolle, F. and Kaehler, J. (2002), The Conditional Efficiency of Signalling: An Experimental Investigation, Europa Universität Viadrina, Mimeo.
Camerer, C. (1988), Gifts as Economic Signals and Social Symbols, *American Journal of Sociology* 94: S180–S214.
Carmichael, H.L. and MacLeod, W.B. (1997), Gift Giving and the Evolution of Cooperation, *International Economic Review* 38: 485–509.
Coleman, J.S. (1984), Introducing Social Structure into Economic Analysis, *American Economic Review (Papers and Proceedings)* 74: 84–88.
Coleman, J.S. (1991), Constructed Organization: First Principles, *Journal of Law, Economics, and Organization* 7: 7–23.
Dasgupta, P. (1988), Trust as a Commodity, in: D. Gambetta (ed.) *Trust: Making and Breaking Cooperative Relations*, Oxford: Blackwell, 49–72.
Gambetta, D. (ed.) (1988), *Trust: Making and Breaking Cooperative Relations*, Oxford: Blackwell.
Krebs, J.R. and Dawkins, R. (1984), Animal Signals: Mind Reading and Manipulation, in: J.R. Krebs and N.B. Davies (eds), *Behavioural Ecology: An Evolutionary Approach*, Oxford: Balckwell, 380–402.
Raz, J. (1975), *Practical Reason and Norms,* vol., Princeton: Princeton University Press.
Smith, A. (1776 [1981]), *An Inquiry into the Nature and Causes of the Wealth of Nations,* vol. I, Indianapolis: Liberty Fund.
Spence, M. (1973), Job Market Signaling, *Quarterly Journal of Economics* 87: 355–374.
Spence, M. (1974), *Market Signaling: Informational Transfer in Hiring and Related Screening Processes,* vol., Cambridge, MA: Harvard University Press.
Williamson, O.E. (1993), Calculativeness, Trust, and Economic Organization, *Journal of Law and Economics* 36: 453–486.
Zahavi, A. (1975), Mate Selection – A Selection for a Handicap, *Journal of Theoretical Biology* 53: 205–214.